

TP 0 — Partie 1: Découverte de l'interface d'Amazon Web Service (AWS)

TP 0 — Partie 1: Découverte de l'interface d'Amazon Web Service (AWS)

Objectifs

1. Création du compte AWS Educate
2. Exploration
- 3 Création d'un espace de stockage Amazon Simple Storage Service (S3)
4. Copie des données dans votre espace de stockage
5. Création d'un IAM, Identity and Access Management
6. Création d'une clef SSH
7. Création d'une machine virtuelle
8. Connexion à sa VM
9. Jouer avec sa VM
10. Eteindre la VM

Objectifs

Ce TP a pour but de vous familiariser avec l'interface d'Amazon Web Service (AWS). Pendant ce TP vous allez :

- Créer un compte AWS educate si ce n'est pas déjà fait
- Copier des données dans votre espace de stockage Amazon Simple Storage Service (S3)
- Lancer une machine virtuelle (VM) et s'y connecter
- Eteindre votre VM

1. Création du compte AWS Educate

Suivez les instructions à partir du mail AWS Educate reçu sur votre adresse ENSAI. Créez vous bien un "starter account". Cela vous permettra d'utiliser la plateforme AWS, mais sans jamais utiliser votre carte de crédit, et donc sans risque de devoir payer quoi que ce soit. Dans le processus de création, vous allez recevoir deux mails. Un premier pour vous demander de confirmer votre mail, un un second (qui arrivera quelques minutes plus tard) qui vous permettra de finaliser la création de votre compte.

Une fois votre compte créé, connectez vous à la plateforme AWS Educate, et cliquez sur "Classroom & Credit" (*note : toutes les images proviennent d'un compte avec l'accès "Educator", il est possible que les interfaces ou que la marche à suivre soient un peu différentes.*)

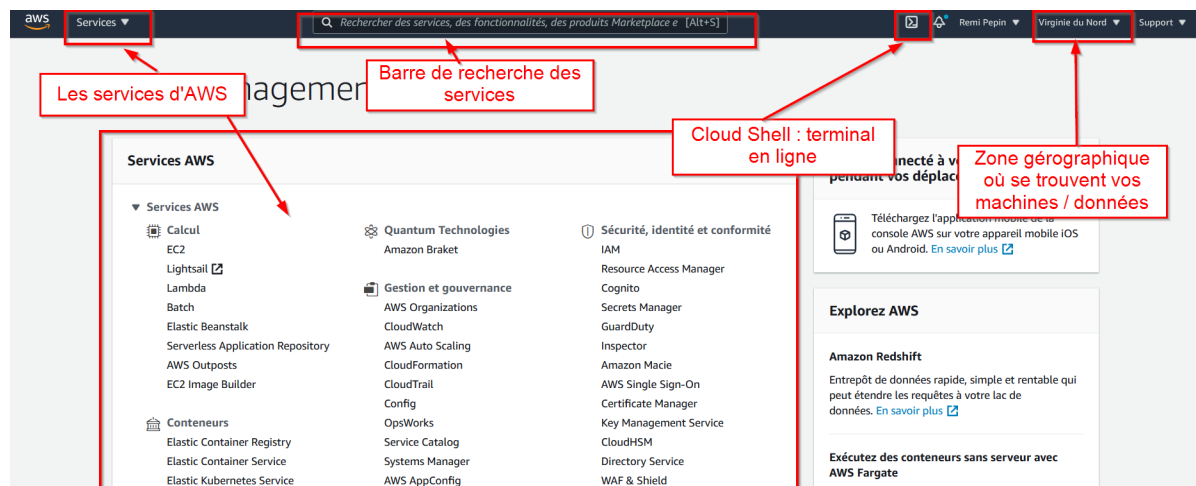
Sur la page suivante cliquez sur "Go to classroom"

Une fenêtre va s'ouvrir, cliquez sur "Continue"

Vous serez ainsi redirigé vers "Vocareum", l'application tierce qui gère les "classrooms". Sélectionnez le bon cours.

Enfin sur la dernière fenêtre, cliquons sur "AWS console". Cela vous redirigera vers la plateforme AWS. Il vous est rappelé d'utiliser le compte AWS qui vous est fourni de manière responsable. C'est à dire de penser à éteindre toutes les instances que vous créez une fois qu'elles ne sont plus utilisées.

Le compte AWS que vous allez utiliser pour les TP est localisé en Virginie du Nord. Ne changez pas cela ! Comme votre compte est à but purement scolaire, vous ne disposez pas de l'intégralité des services de la plateforme (vous n'avez pas accès aux informations de facturation par exemple). De même votre compte est un compte généré par Vocareum, et vous ne pouvez pas y accéder sans passer par ce service. Il est donc inutile de le noter. Pour des connections futures à AWS dans le cadre scolaire, passez toujours par le portail AWS Educate, puis Vocareum.



2. Exploration

Dans l'onglet "Services", trouvez, entre autres :

- EC2, le service de calcul
- S3, le service de stockage
- La section dédiée aux bases de données
- La section dédiée au *machine-learning*
- La section dédiée à l'analyse de données

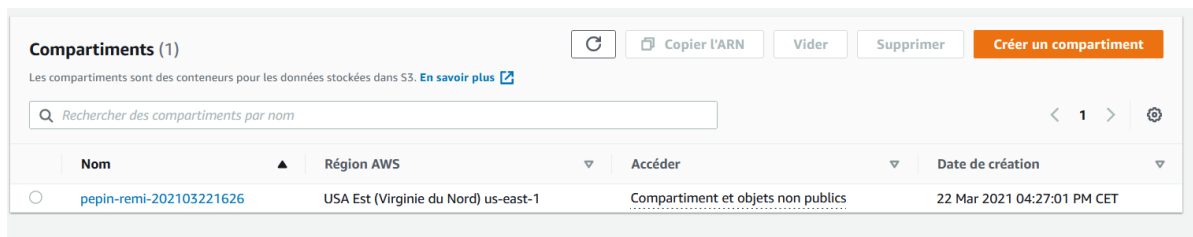
3 Création d'un espace de stockage Amazon Simple Storage Service (S3)

Amazon Simple Storage Service (S3) est la solution de base que propose AWS pour stocker vos données de manière pérenne. Amazon dit assurer une durabilité de vos données de 99,999999999 %. Cela signifie que si vous stockez 10 000 000 objets avec Amazon S3, vous pouvez vous attendre à perdre en moyenne un objet unique une fois tous les 10 000 ans.

Ce stockage est assuré à coût relativement élevé (de l'ordre de ~0,02 \$/Go/mois), sachant que vous payez en sus les opérations de lecture (de l'ordre de ~0,09 \$/Go ; les écritures sont gratuites). 1 To de données vous coûte ainsi 240€ à l'année. Pour comparaison, un disque dur externe SSD d'1 To coûte actuellement ~100€ (pour une durabilité moindre), et un cloud-storage (type dropbox) pour particulier coûte ~10€ / mois pour 2 To (pour une durabilité comparable). S3 est ainsi destiné à des données utilisées régulièrement par d'autres applications hébergées par AWS. D'autres offres de stockage existent comme les archives, pour des données utilisées moins régulièrement, ou les bases de données.

Tous les services que vous propose AWS peuvent nativement lire depuis et écrire vers S3 si vous leur en donnez le droit. Ainsi, les programmes que vous exécutez, les données que vous traitez peuvent être importés/exportés dans S3. Chaque élément hébergé dans S3, appelé "objet", est accessible par une URL **unique**. Vous pouvez restreindre ou au contraire étendre les droits d'accès à vos objets.

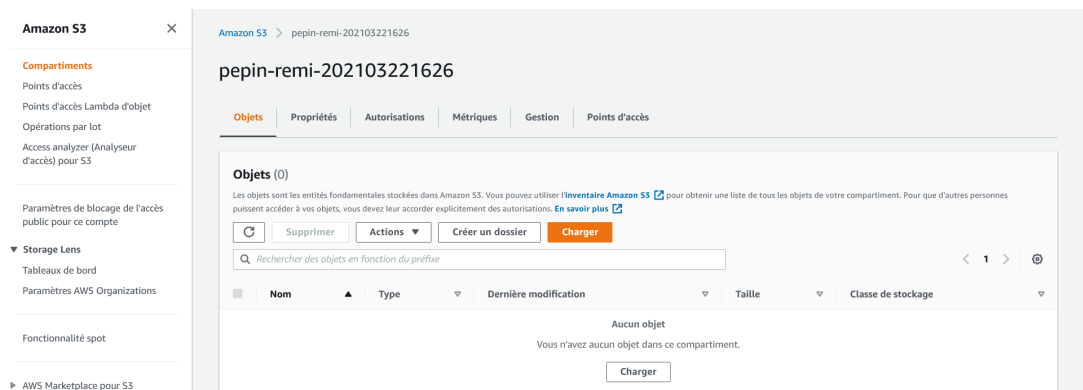
- ☐ Dans la barre de recherche, cherchez "S3" et cliquez dessus



- ☐ Cliquez sur "Créer un compartiment" (en anglais un "bucket")
- ☐ Choisissez un nom unique à votre compartiment (comme votre nom-prénom-et la date et heure du jour)
- ☐ Laissez toutes les valeurs par défaut et descendez en bas de la page pour créer votre compartiment

4. Copie des données dans votre espace de stockage

- ☐ Cliquez sur le nom de votre compartiment pour aller sur sa page



- ☐ À partir du bouton "Charger", ajoutez un fichier (par exemple le fichier json que vous trouverez sur Moodle) à votre compartiment. Laissez toutes les valeurs par défaut.

Charger

Ajoutez les fichiers et dossiers que vous souhaitez charger dans S3. Pour charger un fichier d'une taille supérieure à 160 Go, utilisez la CLI AWS, le kit SDK AWS ou l'API REST Amazon S3. [En savoir plus](#)

Faites glisser et déposez les fichiers et dossiers que vous souhaitez charger ici, ou sélectionnez **Ajouter des fichiers** ou **Ajouter des dossiers**.

Fichiers et dossiers (0)
 Tous les fichiers et dossiers de cette table seront chargés.

	Nom	Dossier	Type	Taille
Aucun fichier ou dossier				
Vous n'avez sélectionné aucun fichier ou dossier à charger.				

- ☐ Une fois le chargement terminé cliquez sur votre fichier. Vous arriverez sur une page similaire avec les lien pour accéder à votre fichier. Aussi bien S3 que HTTP

Présentation de l'objet	
Propriétaire remi.pepin	URI S3 s3://pepin-remi-202103221626/rule20210219-184802.jsonl
Région AWS USA Est (Virginie du Nord) us-east-1	Nom de la ressource Amazon (ARN) arn:aws:s3::pepin-remi-202103221626/rule20210219-184802.jsonl
Dernière modification 22 Mar 2021 04:46:05 PM CET	Balise d'entité (Etag) 433dfec8145842d4a565ec9033be30fa
Taille 904.0 o	URL de l'objet https://pepin-remi-202103221626.s3.amazonaws.com/rule20210219-184802.jsonl
Type jsonl	
Clé rule20210219-184802.jsonl	

5. Création d'un IAM, Identity and Access Management

Une grande partie de la gestion de la sécurité d'AWS se base sur les IAM et les rôles qu'ils possèdent. Actuellement si on ne fait rien il est impossible d'accéder à nos données sur S3 depuis une machine virtuelle (on appelle cette politique de liste blanche. On commente par tout interdire, et on ouvre manuellement les accès). Nous allons remédier à cela en créant un rôle qui permet d'accéder à S3. Bien que les écrans et options sont spécifiques à AWS, cette notion de rôle avec des droits que l'on va appliquer à nos machines est centrale dans la gestion de la sécurité dans le cloud computing

- ☐ Dans la barre de recherche tapez **IAM**
- ☐ Puis allez dans **Rôles** et **Créer un rôle**

Identity and Access Management (IAM)

Tableau de bord

▼ Gestion des accès

Groupes

Utilisateurs

Rôles

Stratégies

Fournisseurs d'identité

Paramètres du compte

▼ Rapports d'accès

Analyseur d'accès

Règles d'archivage

Analyseurs

Paramètres

Rapport sur les informations d'identification

1

2

Créer un rôle Supprimer le rôle

Recherche

Nom du rôle ▼

AWSServiceRoleForAmazonEKS

□ Ensuite sélectionnez Service AWS, EC2 et Suivant

Service AWS
EC2, Lambda et autres

Autre compte AWS
Appartenant à vous ou à un tiers

Identité web
Cognito ou tout fournisseur OpenID

Fédération SAML 2.0
Votre annuaire d'entreprise

Permet aux services AWS d'exécuter des actions en votre nom. [En savoir plus](#)

Choisir un cas d'utilisation

Cas d'utilisation courants

EC2

Allows EC2 instances to call AWS services on your behalf.

Lambda

Allows Lambda functions to call AWS services on your behalf.

Où sélectionner un service pour afficher ses cas d'utilisation

API Gateway	CloudWatch Events	EKS	IoT Things Graph	Redshift
AWS Backup	CodeBuild	EMR	KMS	Rekognition
AWS Chatbot	CodeDeploy	ElastiCache	Kinesis	RoboMaker
AWS Marketplace	CodeGuru	Elastic Beanstalk	Lake Formation	S3
AWS Support	CodeStar Notifications	Elastic Container Registry	Lambda	SMS
Amplify	Comprehend	Elastic Container Service	Lex	SNS
AppStream 2.0	Config	Elastic Transcoder	License Manager	SWF
AppSync	Connect	ElasticLoadBalancing	MQ	SageMaker
Application Auto Scaling	DMS	Forecast	Machine Learning	Security Hub
Application Discovery Service	Data Lifecycle Manager	GameLift	Macie	Service Catalog
Batch	Data Pipeline	Global Accelerator	Managed Blockchain	Step Functions
Budget	Data Brew	Glue	MediaConvert	Storage Gateway

* Obligatoire

Annuler

Suivant : Autorisations

- ☐ Dans la barre de recherche cherchez `s3` puis sélectionnez `AmazonS3FullAccess`. Cela vous donnera tous les droits pour interagir avec votre stockage S3

▼ *Attach stratégies d'autorisation*

Choisissez une ou plusieurs stratégies à attacher à votre nouveau rôle.

Créer une stratégie

Stratégies de filtre 8 résultats affichés

	Nom de la stratégie	Utilisé comme
<input type="checkbox"/>	AmazonDMSRedshiftS3Role	Aucun
<input checked="" type="checkbox"/>	AmazonS3FullAccess	Aucun
<input type="checkbox"/>	AmazonS3OutpostsFullAccess	Aucun
<input type="checkbox"/>	AmazonS3OutpostsReadOnlyAccess	Aucun
<input type="checkbox"/>	AmazonS3ReadOnlyAccess	Permissions policy (2)
<input type="checkbox"/>	IVSRecordToS3	Aucun
<input type="checkbox"/>	QuickSightAccessForS3StorageManagementAnalyticsReadOnly	Aucun
<input type="checkbox"/>	S3StorageLensServiceRolePolicy	Aucun

► Définir une limite d'autorisations

* Obligatoire

Annuler Précédent Suivant : Balises

- ☐ Ne définissez aucune balise et appuyez sur suivant
- ☐ Enfin donnez un nom à votre rôle, comme `Ec2_s3_fullAccess` puis créez le rôle.

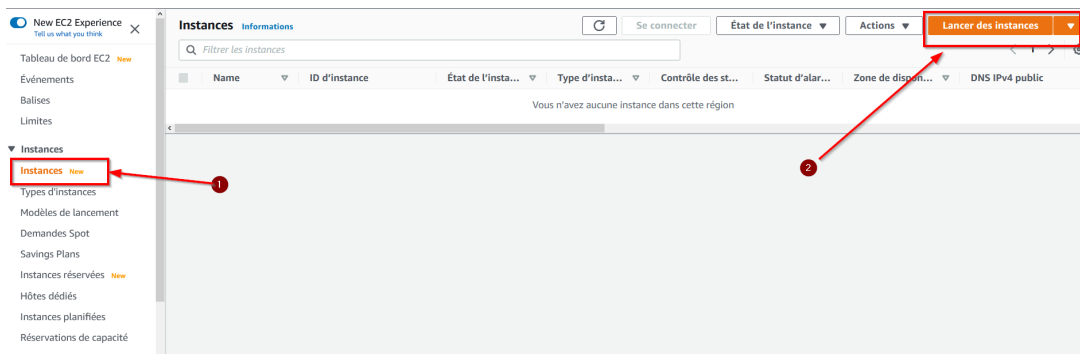
6. Création d'une clef SSH

SSH (Secure SHell) permet de se connecter de façon sécurisée à un système Unix comme le cluster de l'Ensaï qui est sous une distribution Linux. Pour plus d'information, je vous conseille de lire le début de cette [page web](#)

- ☐ Dans la barre de recherche, cherchez "paire de clés" et cliquez dessus
- ☐ Cliquez sur "Créer une paire de clés"
- ☐ Donnez lui un nom (par ex: "ensai_big_data_TP"), sélectionnez le format PPK si vous utilisez une machine windows, et pem si vous utilisez une machine sous Linux / Mac Os, et cliquez sur "créer"
- ☐ Cela va lancer le téléchargement d'un fichier, le perdez pas !

7. Création d'une machine virtuelle

- ☐ Dans la barre de recherche, cherchez "EC2" et cliquez dessus



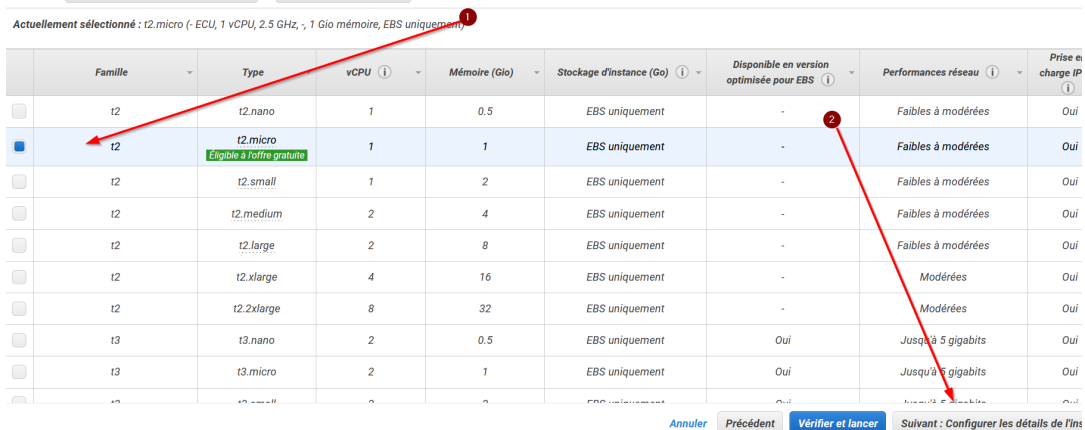
☐ Cliquez sur **Lancer une instance**

☐ Vous devez choisir l'image de la machine à créer appelé ici AMI pour *Amazon Machine Image*. Une image contient notamment le système d'exploitation. Choisissez la première: Amazon Linux 2 AMI.



☐ Vous choisissez ensuite la configuration de VM. Par exemple, vous pouvez choisir une machine d'usage général à 1 cœurs **t2.micro** pour avoir une machine de faible puissance mais peu chère (0.012\$/heure) pour une machine plus puissante comme une **t2.xlarge** (0.188\$/heure). Comme la facturation est au temps d'utilisation, pensez à éteindre vos machines à la fin du TP !

Étape 2 : Choisir un type d'instance



- ☐ Sur l'écran suivant sélectionnez votre IAM : **Configurer les détails de l'instance** : une instance en fonction de vos besoins. Vous pouvez lancer plusieurs instances à partir de la même AMI, demander des instances Spot pour des cas encore.

Nombre d'instances [Lancer dans le groupe Auto Scaling](#)

Option d'achat ☐ Demander des instances Spot

Réseau [Créer un nouveau VPC](#)

Sous-réseau [Créer un nouveau sous-réseau \(subnet\)](#)

Attribuer automatiquement l'adresse IP publique

Groupe de placement ☐ Ajoutez une instance au groupe de placement.

Réserve de capacité

Fournisseur de jonction de domaines [Créer un nouveau répertoire](#)

Rôle IAM [Créer un nouveau rôle IAM](#)

- ☐ Validez cet étape et les 2 suivantes (stockage et balise)
- ☐ Sur l'écran des groupes de sécurité ajoutez une règle SSH avec la source 18.206.107.24/29

Description: launch-wizard-4 created 2021-03-23T11:57:11.987+01:00

Type	Protocole	Plage de ports	Source	Description
SSH	TCP	22	Personnalisés 0.0.0.0/0	par exemple SSH
SSH	TCP	22	Personnalisés 18.206.107.24/29	par exemple SSH

[Ajouter une règle](#)

Ceci va rendre possible l'accès à votre machine via un *cloud shell* (un shell dans votre navigateur). Cette configuration est spécifique à AWS. Elle se trouve dans la [documentation officielle](#)

- ☐ Ensuite validez cette étape et la suivante
- ☐ Choisissez la bonne paire de clés
- ☐ Et voilà! Votre VM est en cours de lancement. Cliquez sur **Affichez les instances** et attendez quelques minutes !

8. Connexion à sa VM

- ☐ Une fois l'instance lancée vous pouvez accéder à son écran d'administration en cliquant sur son **id d'instance**

EC2 > Instances > i-02e3a617255cd46af

Résumé de l'instance pour i-02e3a617255cd46af Informations

Mis à jour il y a moins d'une minute

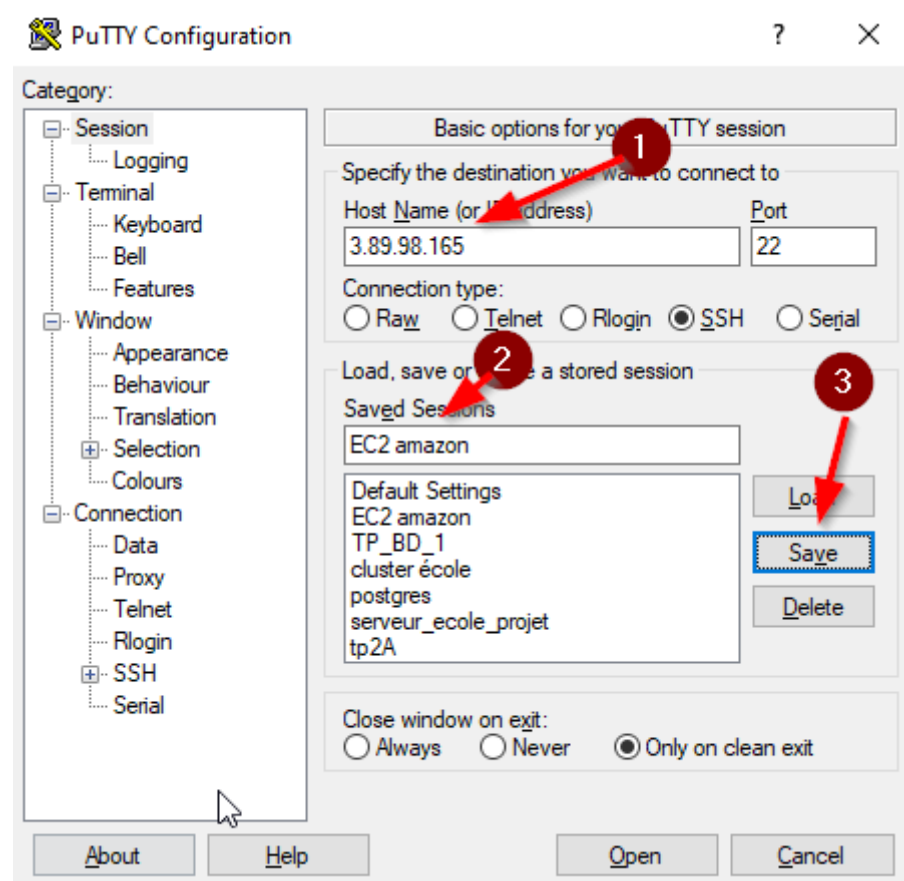
ID d'instance i-02e3a617255cd46af	Adresse IPv4 publique 3.89.98.165 adresse ouverte	Adresses IPv4 privées 172.31.85.171
État de l'instance En cours d'exécution	DNS IPv4 public ec2-3-89-98-165.compute-1.amazonaws.com adresse ouverte	DNS IPv4 privé ip-172-31-85-171.ec2.internal
Type d'instance t2.xlarge	Adresses IP Elastic -	ID de VPC vpc-e2567a98
Recherche d'AWS Compute Optimizer Inscrivez-vous à AWS Compute Optimizer pour obtenir des recommandations. En savoir plus	Rôle IAM -	ID de sous-réseau subnet-5db4ab75

Détails Sécurité Mise en réseau Stockage Vérifications de statut Surveillance Balises

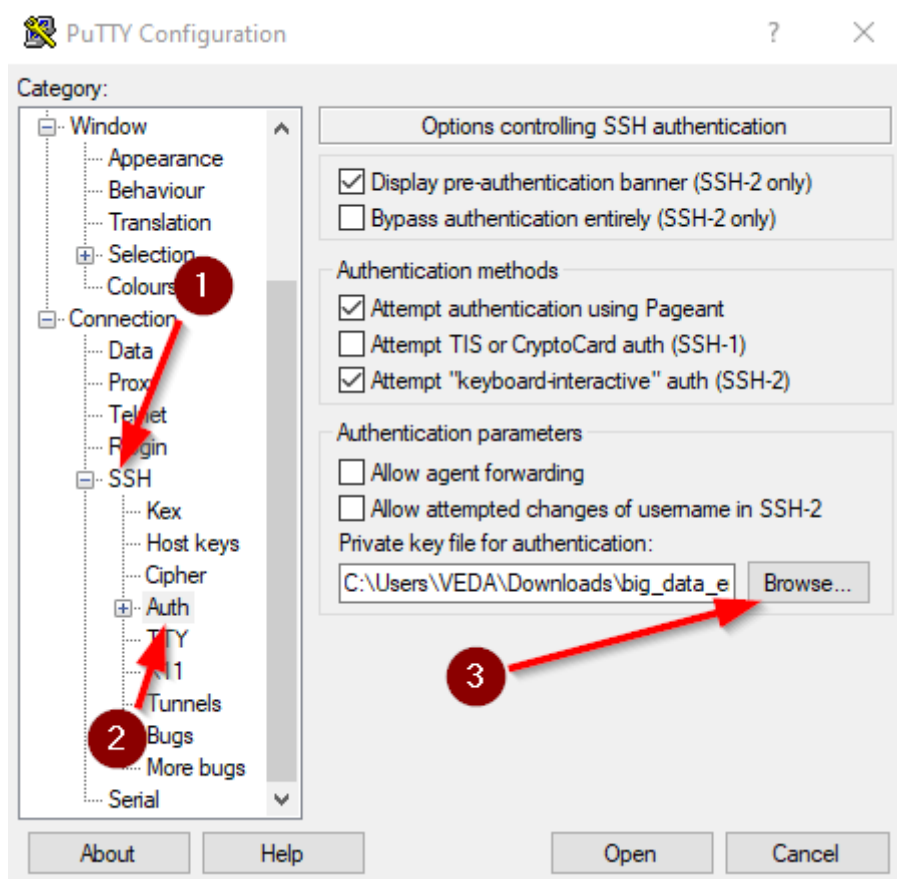
▼ Détails de l'instance Informations

Vous y trouverez de nombreuses informations, mais surtout l'adresse **IPv4 publique** qui est l'adresse IP (Internet Protocol) de votre machine pour y accéder en étant à l'extérieur de la plateforme AWS, par exemple depuis votre ordinateur.

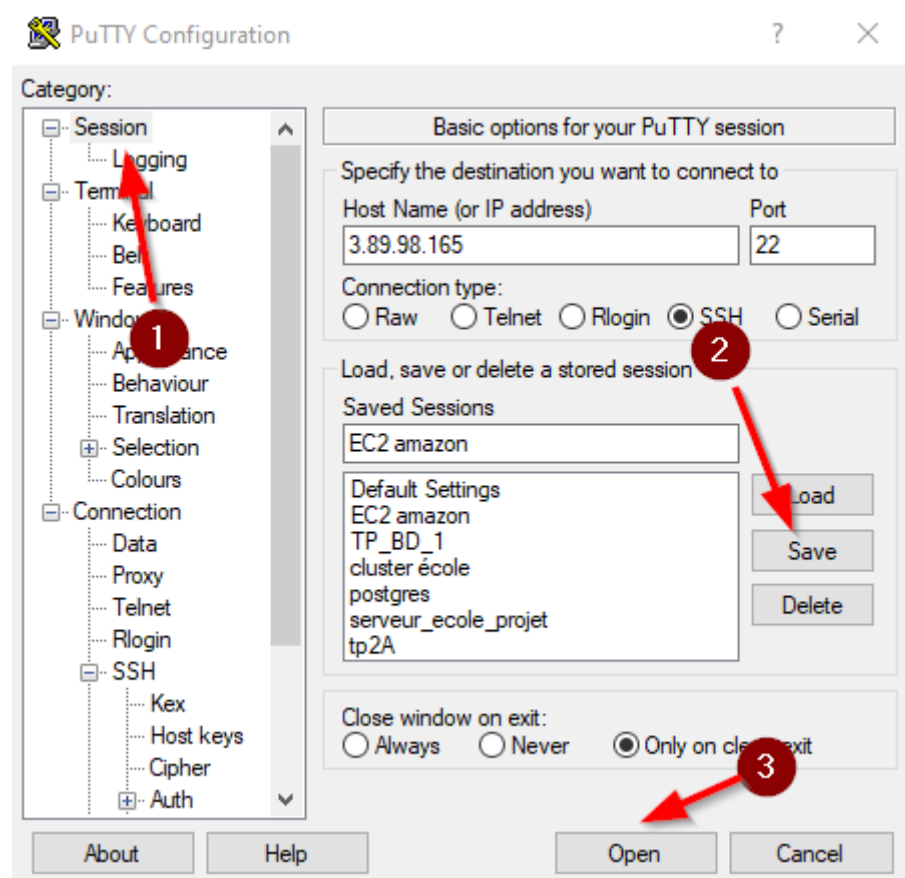
- ☐ Lancez PuTTY
- ☐ Dans la partie **Host Name** saisissez l'adresse publique de votre serveur
- ☐ Dans la partie **Saved Session** rentrez le nom que vous voulez puis cliquez sur **Save**



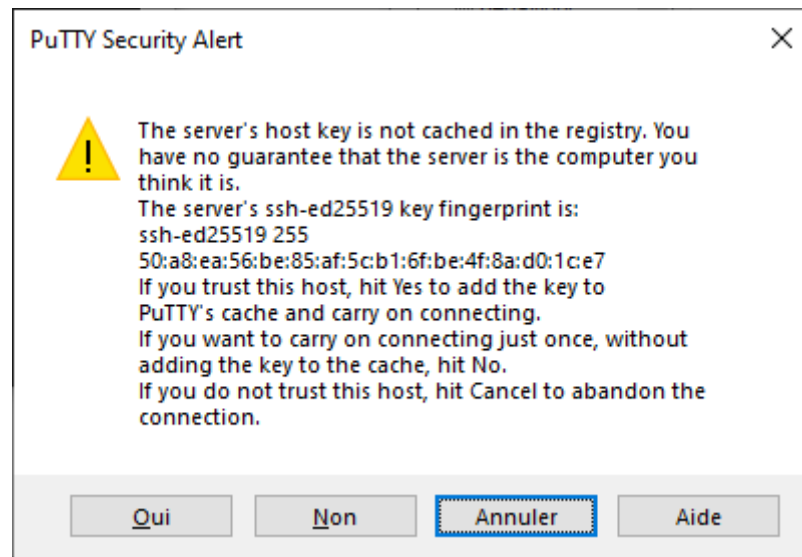
- ☐ Puis allez dans le menu **SSH** et chargez votre fichier .ppk



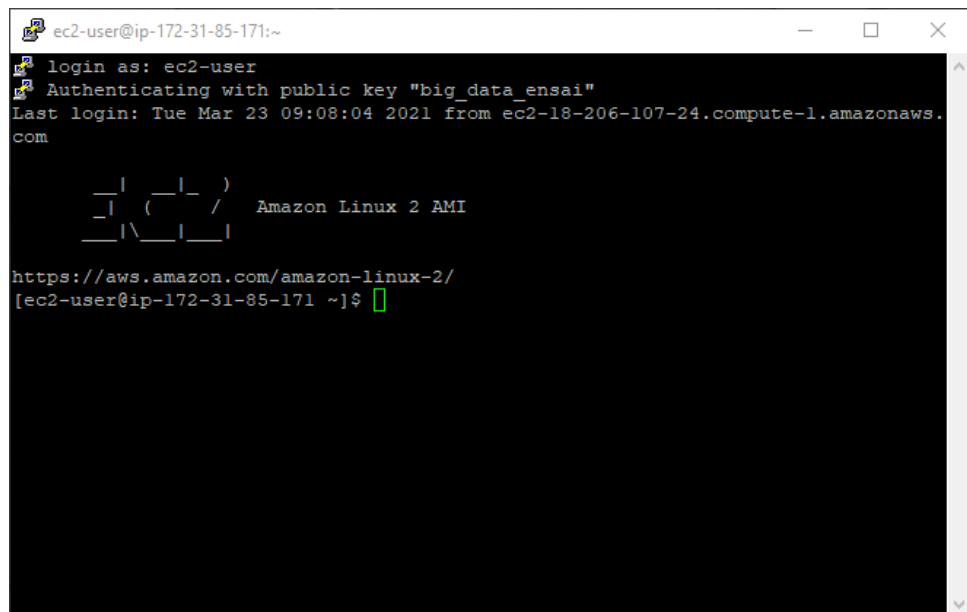
- ☐ Enfin retournez dans l'écran initial, cliquez sur le nom de la session puis sur **save** pour sauvegarder votre configuration. Lancez la session SSH.



- ☐ Une fenêtre semblable va s'ouvrir pour vous prévenir que c'est la première fois que vous vous connectez à cette machine et si vous lui faites confiance. Vous allez cliquer sur **oui**



- ☐ Un terminal va s'ouvrir avec écrit `login as: .` Saisissez `ec2-user` puis validez ([documentation officielle](#))



- ☐ Voilà vous venez de vous connecter à votre machine virtuelle. **Bien que visuellement le terminal se trouve sur votre écran, tout ce que vous allez exécuter dans se terminal sera réalisé sur une machine distante.** Vous pouvez ainsi réaliser de calculs très longs et nécessitant une grande puissance de calcul sur une machine puissance depuis votre ordinateur. Par contre cette machine n'a pas d'interface graphique (GUI : *graphical user interface*) et va nécessiter de connaître quelques rudiments de *bash*.

9. Jouer avec sa VM

- ☐ Lancez python avec la commande `python`. Cela ouvre un *shell* python, similaire à la console IPython de Spyder. Vous pouvez essayer ce petit code

```

1 def fibo(n):
2     if n==1 or n==0:
3         return 1
4     else :
5         return fibo(n-1) + fibo(n-2)
6
7
8 fibo(5)

```

Pour quitter le *shell* python appuyez sur `ctrl+D`

Si le *shell* python n'a pas que peu d'intérêt, cela signifie que vous avez un interpréteur python sur cette machine. Ainsi vous pouvez lancer un code python en faisant `python votre_fichier.py`

- ☐ R n'est pas par défaut installé. Vous pouvez l'installer en saisissant la ligne de commande suivante : `sudo amazon-linux-extras install R4`. Quand l'installation vous demandera si vous voulez vraiment installer ce package appuyez sur `y`. Le terminal va se remplir de texte pendant quelques minutes n'y prêtez pas attention, c'est juste la machine qui vous dit ce qu'elle fait. Ensuite pour lancer R vous devez seulement saisir `R` dans votre terminal.
- ☐ Vous pouvez également télécharger des fichier stockez sur S3. Pour ce faire vous allez saisir la ligne suivante :
`aws s3 cp uri_s3_de_votre_fichier .` (il y a bien un point à la fin de la commande) en remplaçant `uri_s3_de_votre_fichier` par la valeur contenu dans URI s3 sur l'écran de votre fichier sur S3

Présentation de l'objet	
Propriétaire remi.pepin Région AWS USA Est (Virginie du Nord) us-east-1 Dernière modification 22 Mar 2021 04:46:05 PM CET Taille 904.0 o Type jsonl Clé rule20210219-184802.jsonl	URI S3 s3://pepin-remi-202103221626/rule20210219-184802.jsonl Nom de la ressource Amazon (ARN) arn:aws:s3::pepin-remi-202103221626/rule20210219-184802.jsonl Balise d'entité (Etag) 433dfec8145842d4a565ec9033be30fa URL de l'objet https://pepin-remi-202103221626.s3.amazonaws.com/rule20210219-184802.jsonl

Cela devrait produire un résultat similaire à celui-ci

```

1 [ec2-user@ip-172-31-82-150 ~]$ aws s3 cp s3://pepin-remi-
2 202103221626/rule20210219-184802.jsonl .
download: s3://pepin-remi-202103221626/rule20210219-184802.jsonl to
./rule20210219-184802.jsonl

```

- ☐ Fermez votre terminal et retournez sur la page de votre instance EC2. Nous allons maintenant nous y connecter via un *cloud shell*. Pour ce faire cliquez sur `Se connecter`

Résumé de l'instance pour i-0ba0dca1d2fd99bee Informations <div> <div>Se connecter</div> <div>État de l'instance</div> </div>		
ID d'instance i-0ba0dca1d2fd99bee État de l'instance En cours d'exécution	Adresse IPv4 publique 35.174.167.162 adresse ouverte DNS IPv4 public ec2-35-174-167-162.compute-	Adresses IPv4 privées 172.31.82.150 DNS IPv4 privé ip-172-31-82-150.ec2.internal

Vous allez arriver sur une page similaire à celle ci-dessous. Cliquez sur `Se connecter`

EC2 > Instances > i-0ba0dca1d2fd99bee > Se connecter à l'instance

Se connecter à l'instance [Informations](#)

Se connecter à votre instance à i-0ba0dca1d2fd99bee l'aide de l'une de ces options

Connexion d'instance EC2

Session Manager

Client SSH

ID d'instance

i-0ba0dca1d2fd99bee

Adresse IP publique

35.174.167.162

Nom utilisateur

ec2-user

Connectez-vous à l'aide d'un nom d'utilisateur personnalisé ou du nom d'utilisateur par défaut ec2-user de l'AMI utilisée pour lancer l'instance.

Remarque : Dans la plupart des cas, le nom d'utilisateur deviné est correct. Cependant, lisez les instructions d'utilisation de l'AMI pour vérifier si le propriétaire de l'AMI a modifié le nom d'utilisateur par défaut.

Annuler

Se connecter

Après quelques instants vous allez arriver sur votre *cloud shell*

```
Last login: Tue Mar 23 11:17:26 2021 from ec2-18-206-107-24.compute-1.amazonaws.com

 _ _ | _ _ )
 _ | ( _ _ /  Amazon Linux 2 AMI
 _ | \ _ _ | _ _ |

https://aws.amazon.com/amazon-linux-2/
[ec2-user@ip-172-31-82-150 ~]$
```

i-0ba0dca1d2fd99bee
Public IPs: 35.174.167.162 Private IPs: 172.31.82.150

Depuis cette écran vous êtes connecté à votre machine distante. Par exemple tapez la commande suivante `ls` (*list* permet de lister les fichiers dans un dossier). Cela va afficher le nom de votre fichier. Ensuite faites `cat le_nom_de_votre_fichier` (vous avez de autocomplétions avec la touche tabulation; `cat` signifie *concatenate*) pour afficher votre fichier.

10. Eteindre la VM

Le coût d'une VM est fonction de son temps d'utilisation, pas du travail qu'il accomplit. Ainsi, une fois le travail effectué, vous devez éteindre vos VMs ! **Même si le coût horaire est bas, faire tourner une machine EC2 pendant 1 semaine se chiffre en dizaines d'euros!**

Pour éteindre votre VM, allez sur la page d'accueil `EC2 > Instances` en cours d'exécution ou sur la bar de navigation `Instances > Instances`, enfin `Etat de l'instance`. Selon le type d'instance, vous pouvez l'arrêter (**EN:** *stop*, pour la réutiliser plus tard), ou la résilier (**EN:** *terminate*, i.e. la supprimer). Dans les deux cas, les données en mémoire et le stockage local sont perdus, mais dans le premier cas, la configuration (URL et IP) sont conservés.

New EC2 Experience
Tell us what you think

Tableau de bord EC2 New

Événements

Balises

Limites

▼ Instances

Instances New

Types d'instances

Modèles de lancement

Demandes Spot

Savings Plans

Arrêt réussi i-0ba0dca1d2fd99bee

Instances (1/3) Informations

Se connecter

État de l'instance ▲

Actions ▼

Lancer

Filter les instances

	Name	ID d'instance	État de l'insta...	Type d'insta
<input type="checkbox"/>	-	i-02e3a617255cd46af	Arrêté(e)	t2.large
<input type="checkbox"/>	-	i-048005c7bcd927adc	Arrêté(e)	t2.micro
<input checked="" type="checkbox"/>	-	i-0ba0dca1d2fd99bee	En cours d'ex...	t2.micro

Arrêter l'instance

Démarrer l'instance

Redémarrer l'instance

Mettre l'instance en veille prolongée

Résilier l'instance

4/4 vérifications n... Aucune at...

1

2

3